The machine intelligence of images

MVA 2023

The machine intelligence of images – Course Overview

Lectures :

- Introduction to cameras and the industrial ecosystem (F. Guichard / W. Hauser)
- DSLR vs Smartphone – the sensitivity and denoising challenge (F. Guichard / W. Hauser)
- Image Quality Analysis from RAW to Final Image quality (C. Greco / B. Pochon)
- **Color and Dynamic Range (W. Hauser / B. Pochon)**
- Image processing/machine learning implementation constraints (M. Karpushin / W. Hauser)
- Multi-Images / Multi-Sensors (B. Neveu)
- IA from cameras to IQ evaluation (S. Ferradans / B. Pochon)

Participants :

Frederic Guichard
DXOMARK

Claudio Greco
DXOMARK

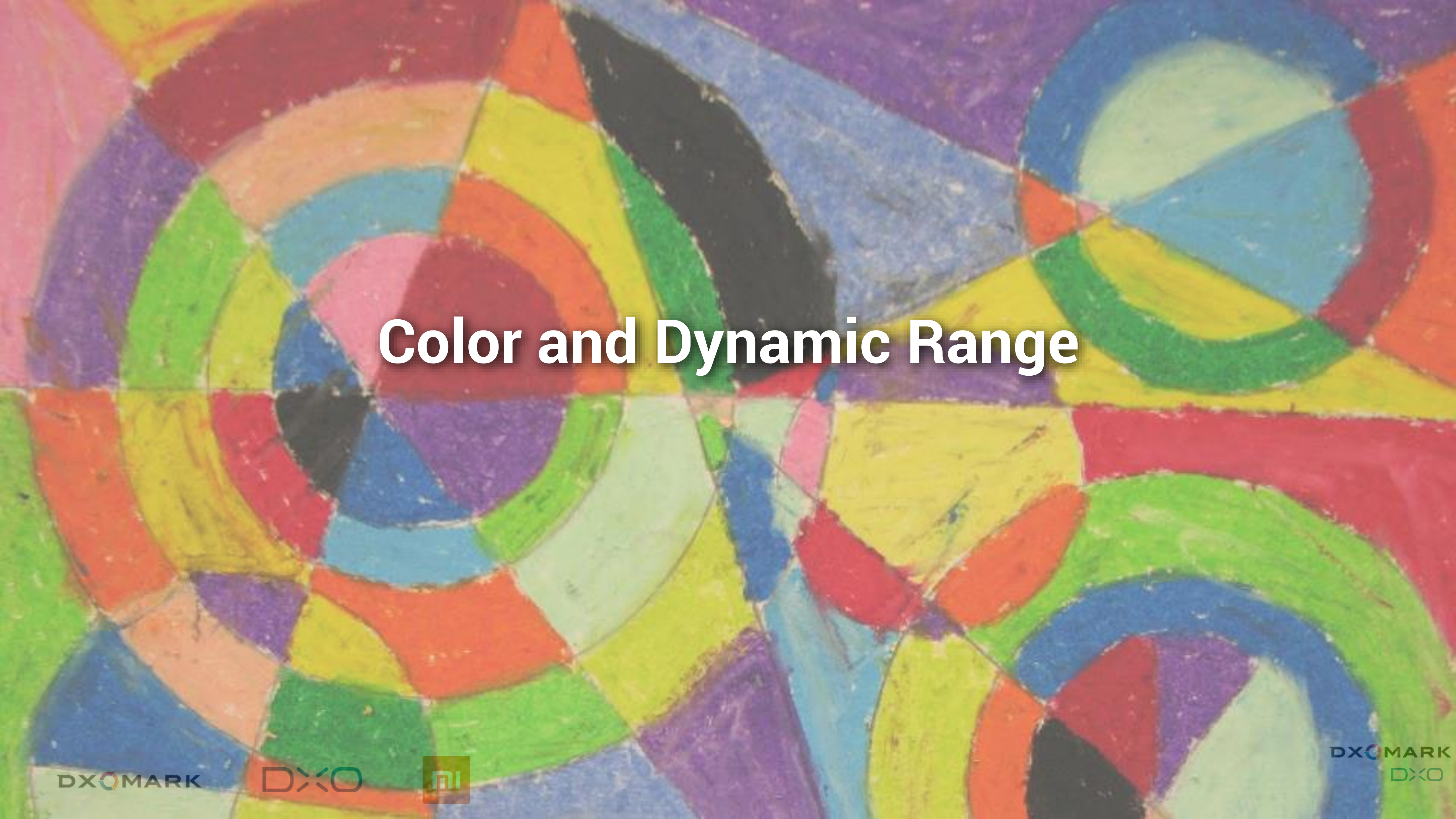
Benoit Pochon
DXOMARK

Sira Ferradans
DXOMARK

Wolf Hauser
DxO

Balthazar Neveu
Xiaomi

Maxim Karpushin
Xiaomi

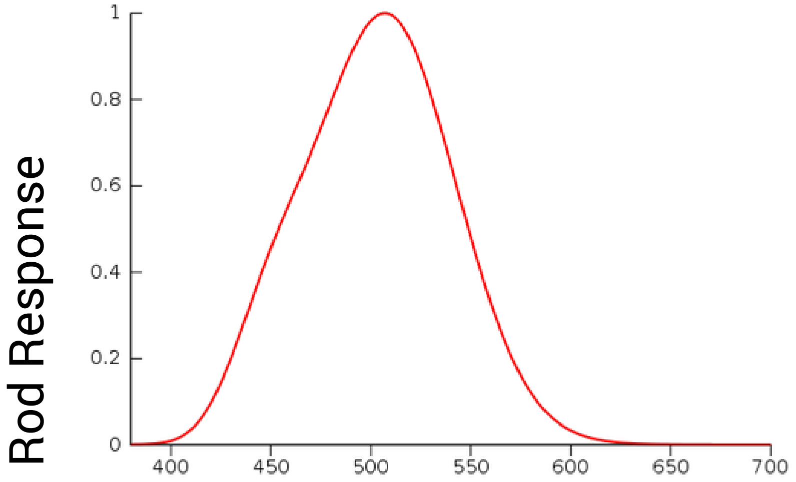


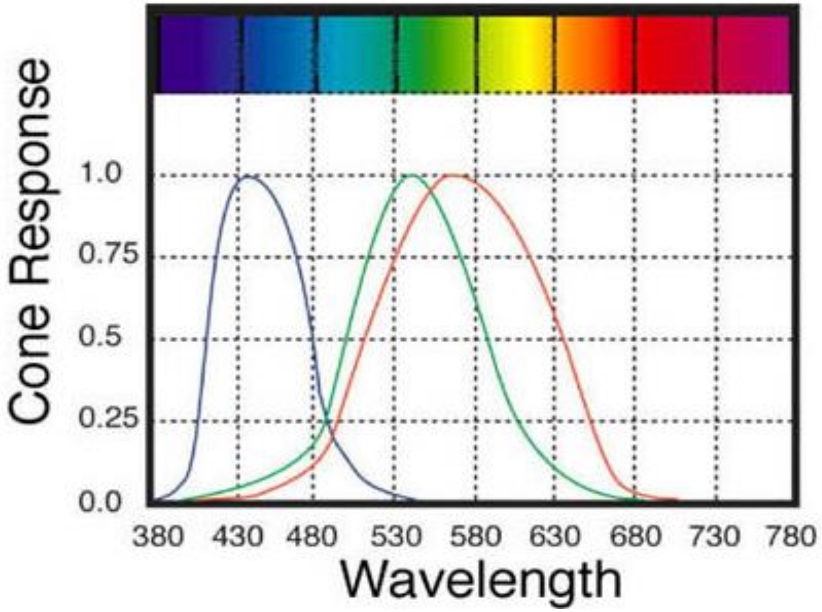
Color and Dynamic Range

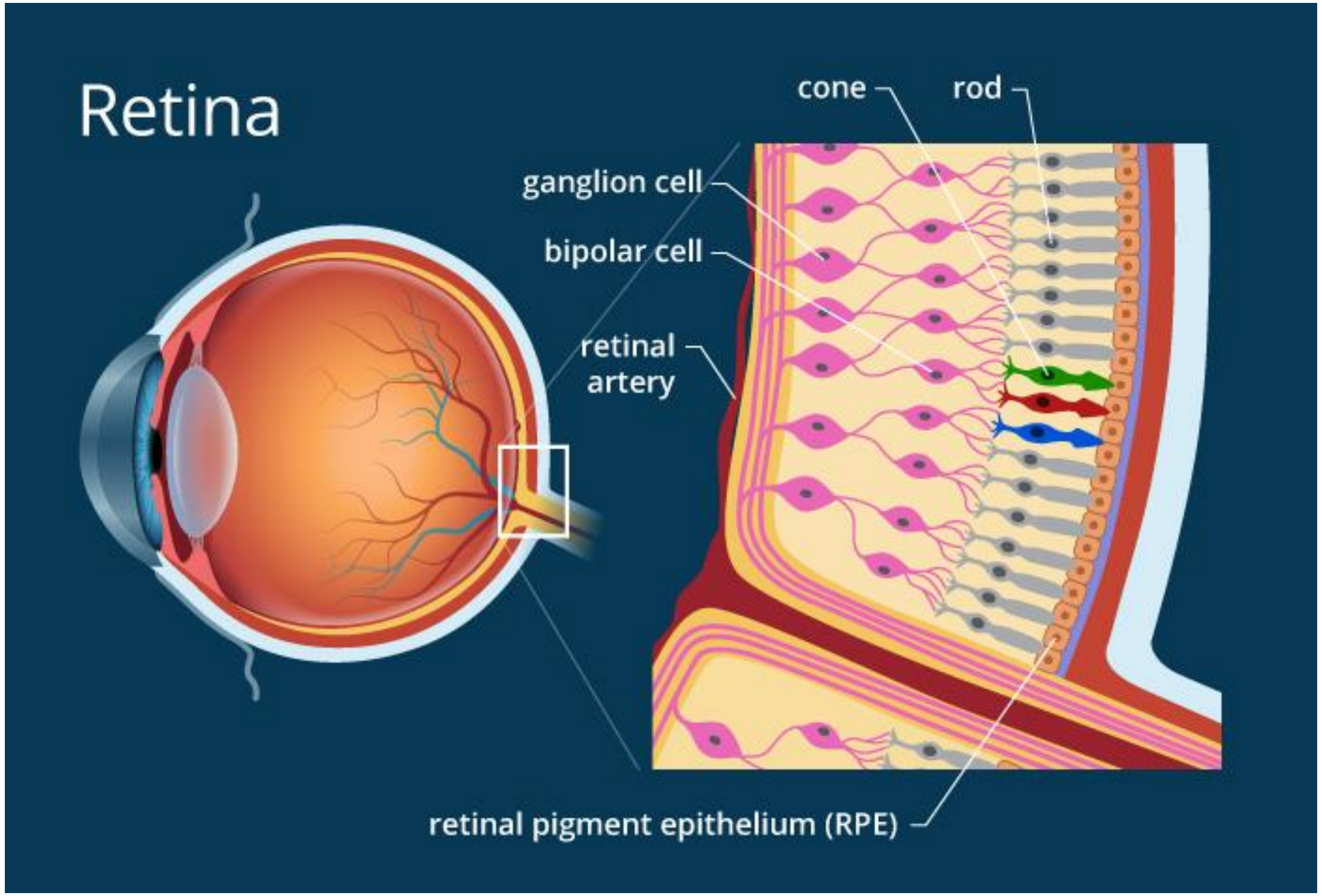
Color and Dynamic Range

- **Human Visual System**
- CIE Color Models
- Reproducing Color
- Color Processing in Cameras

Color is not an object property, but a (human) perception



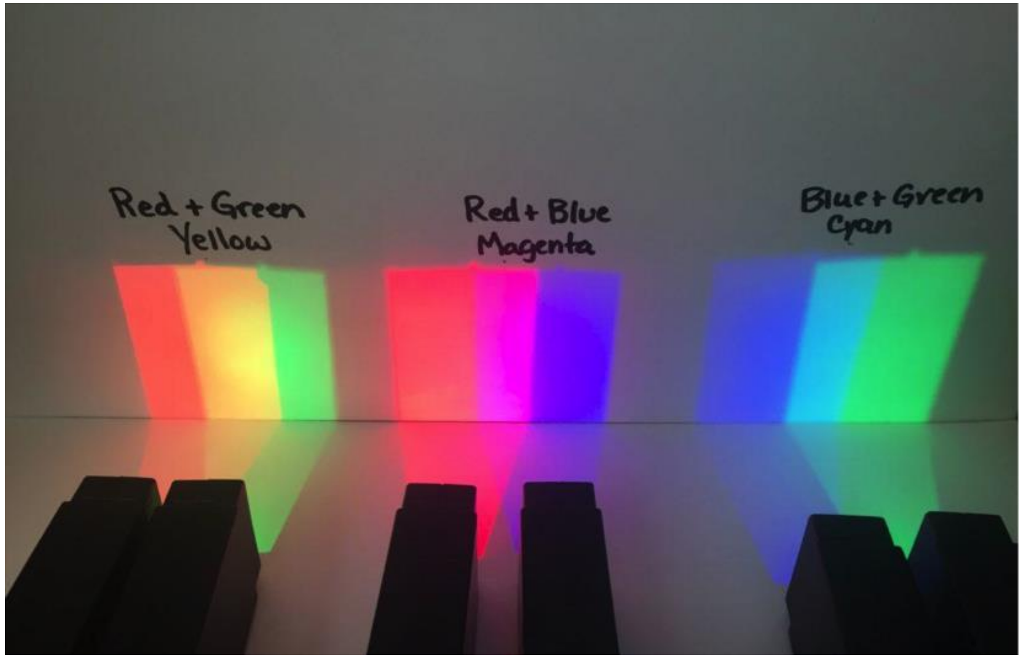




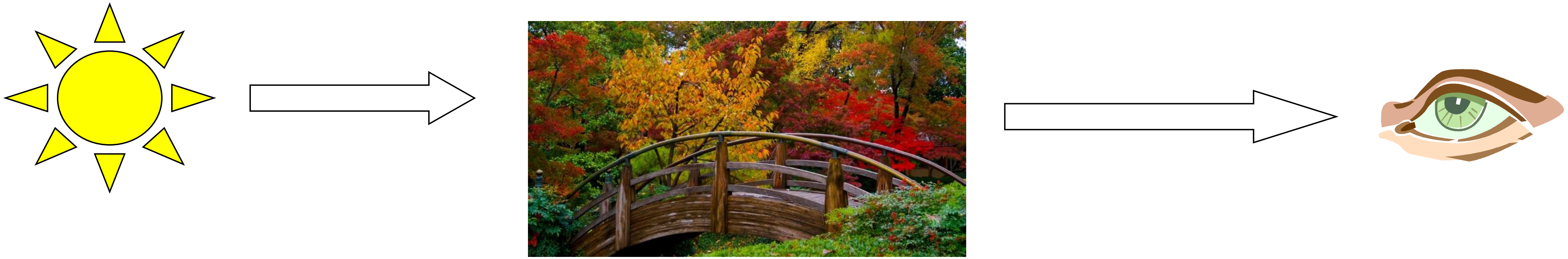
Trichromacy

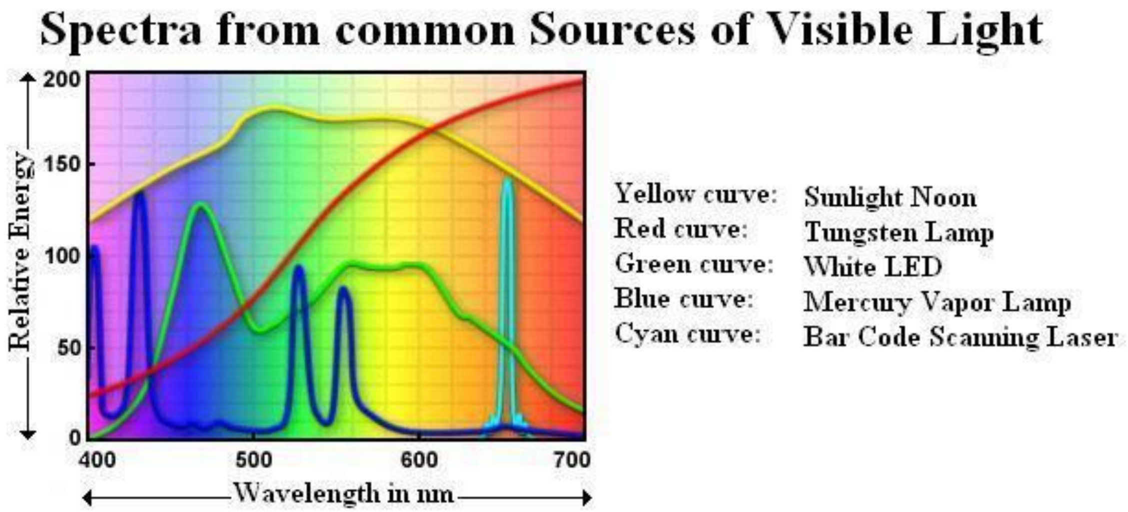
$$L = \int_{\lambda=380nm}^{780nm} \phi(\lambda).\ell(\lambda)d\lambda \qquad M = \int_{\lambda=380nm}^{780nm} \phi(\lambda).m(\lambda)d\lambda \qquad S = \int_{\lambda=380nm}^{780nm} \phi(\lambda).s(\lambda)d\lambda$$

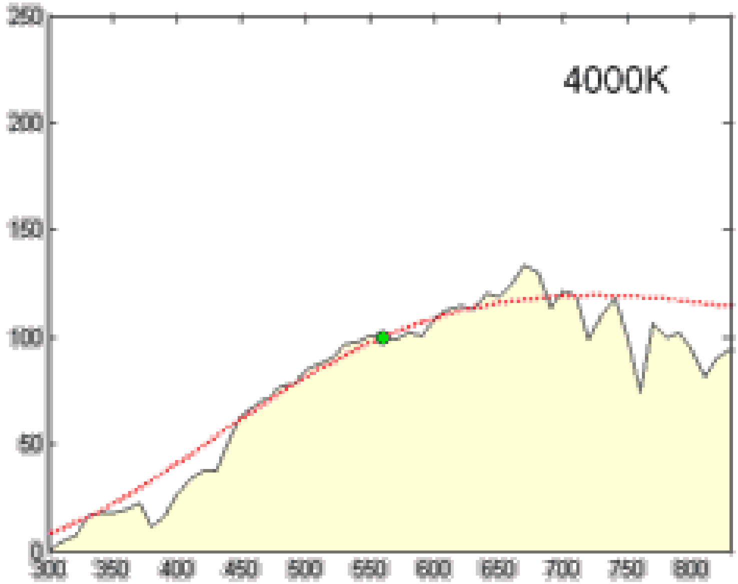
Primary colors: Any color can be obtained as additive mix of 3 independent colors



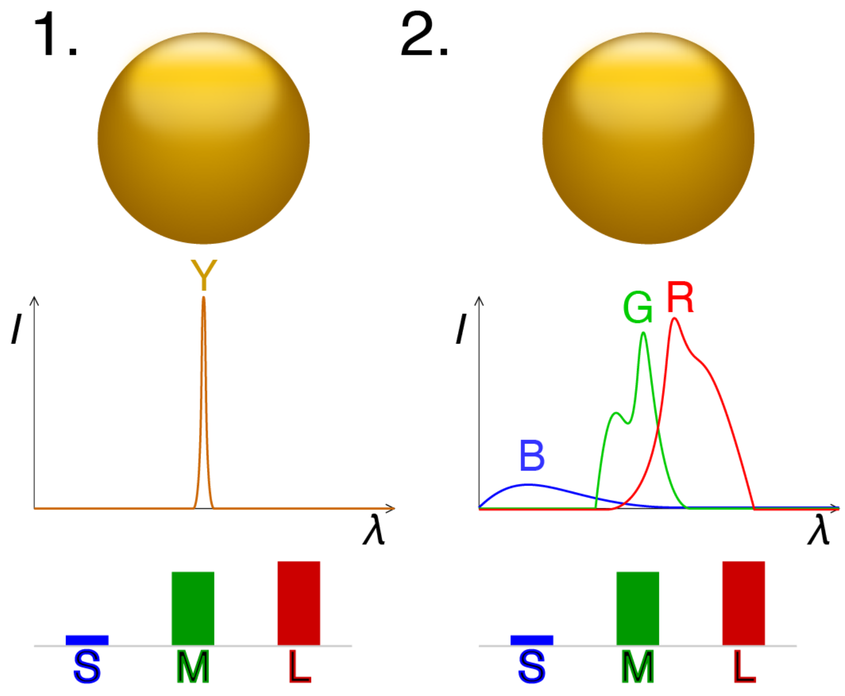
Light source





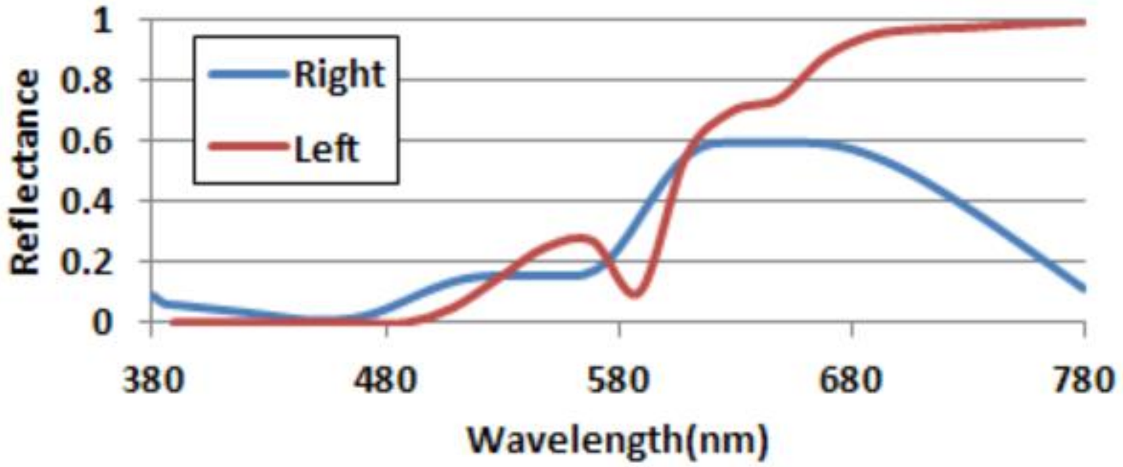


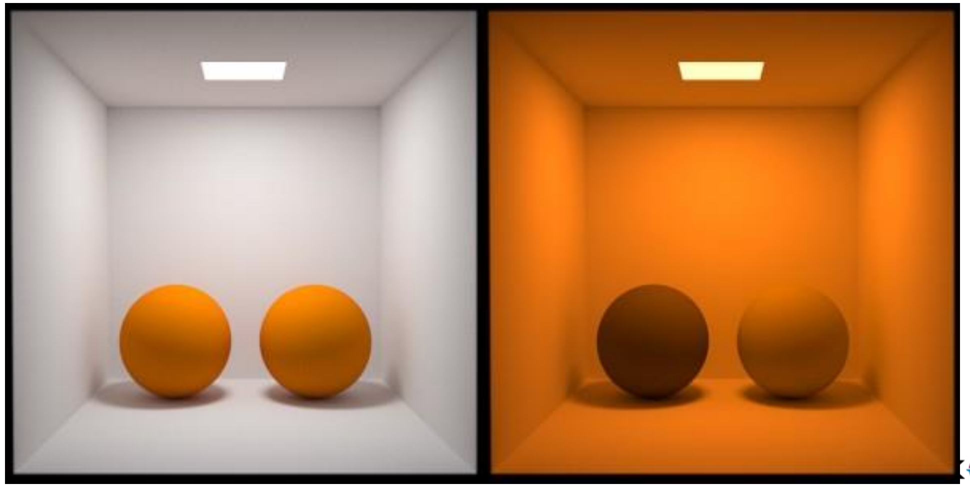
Metamerism



Perceived matching of colors with different (nonmatching) spectral power distributions

Surfaces are usually metamers under one specific illuminant

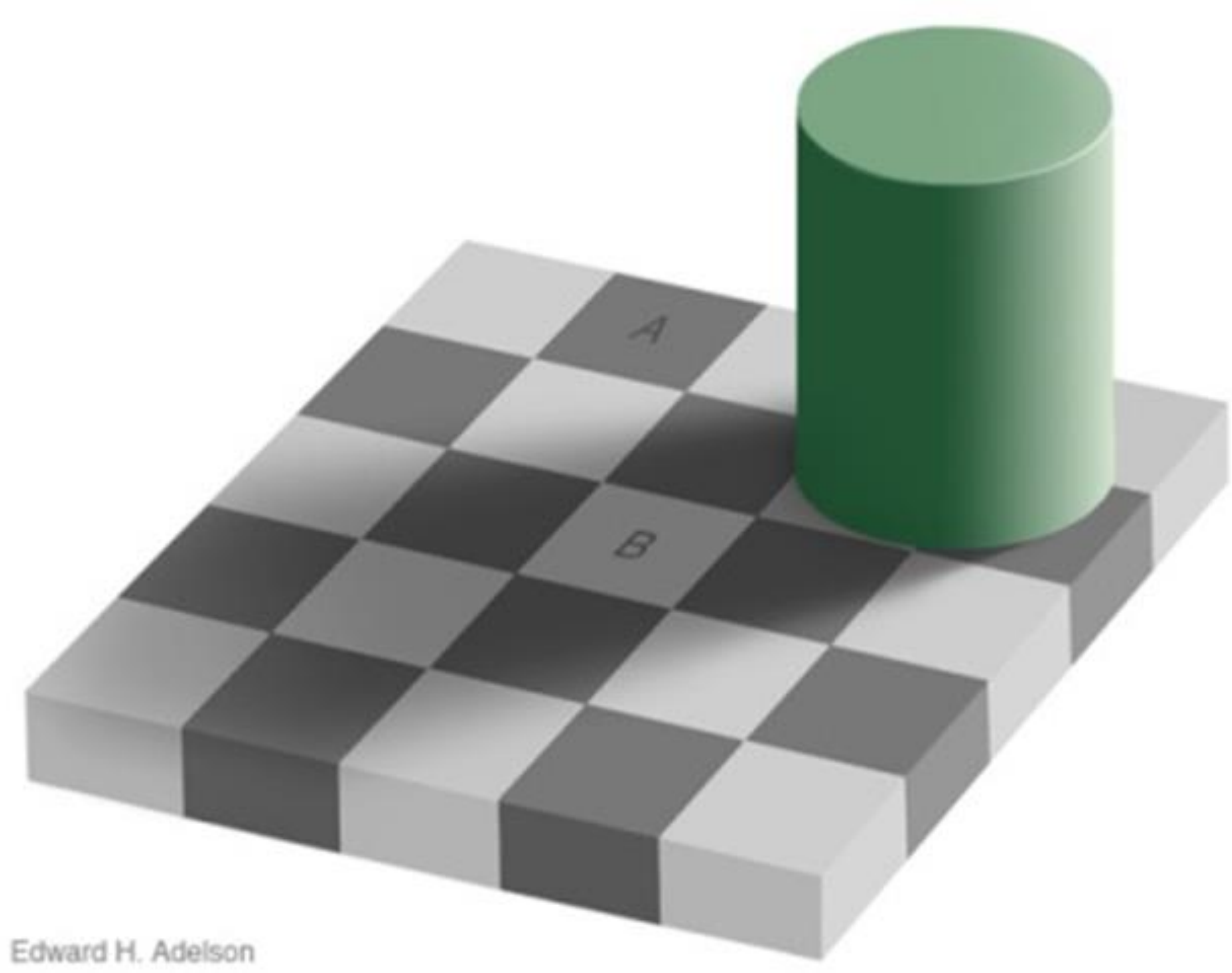




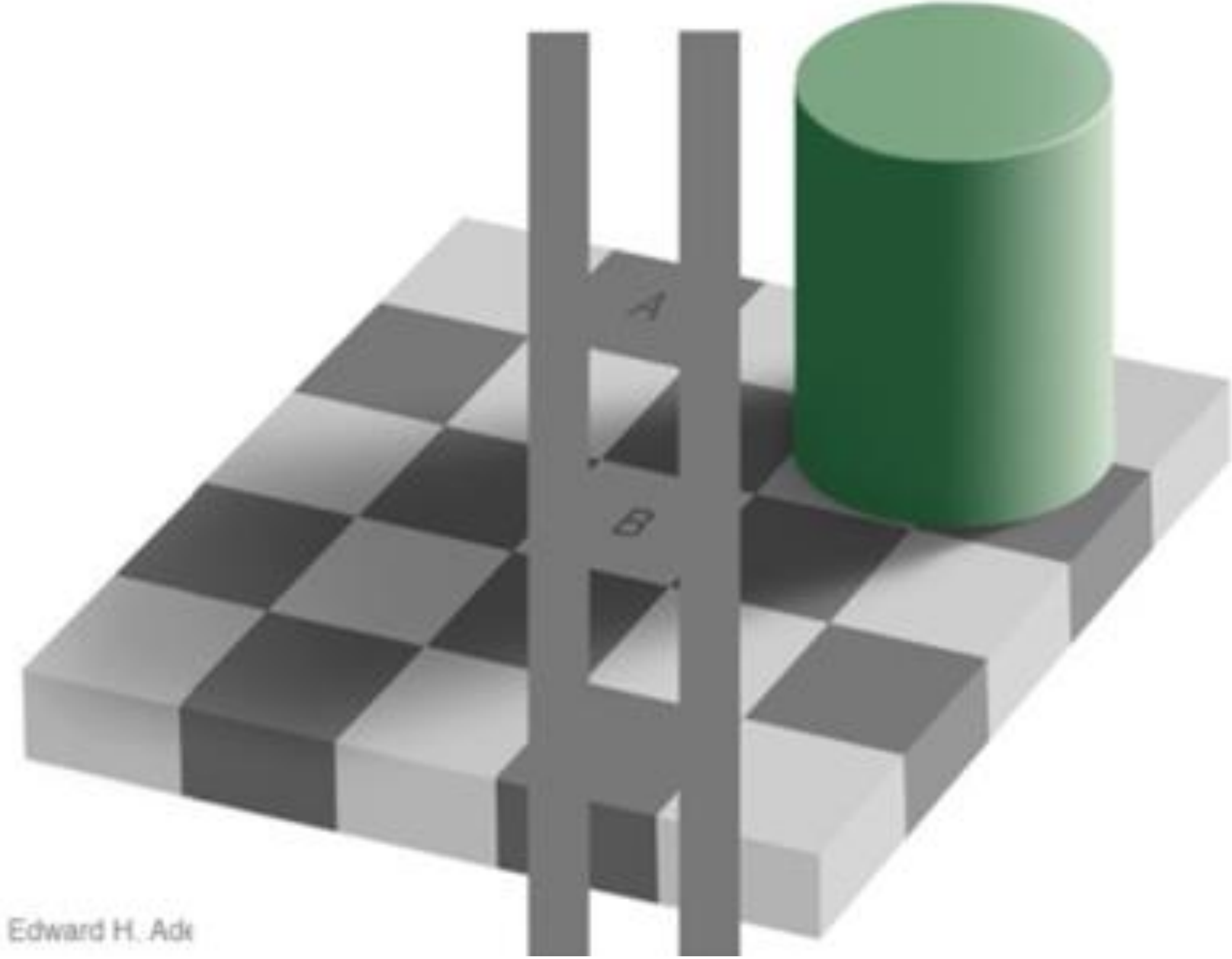
D65

Sodium vapor

Adaptation mechanisms

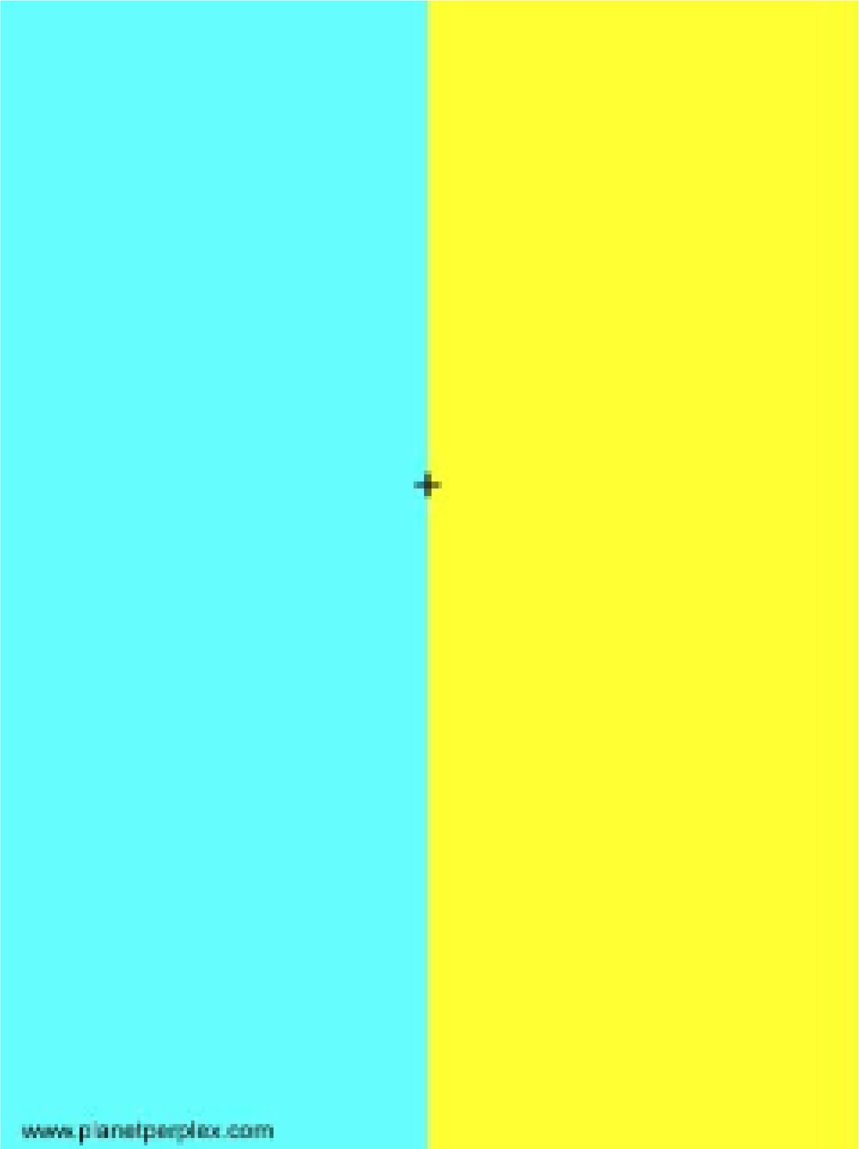


Adaptation mechanisms



Adaptation mechanisms

Adaptation mechanisms

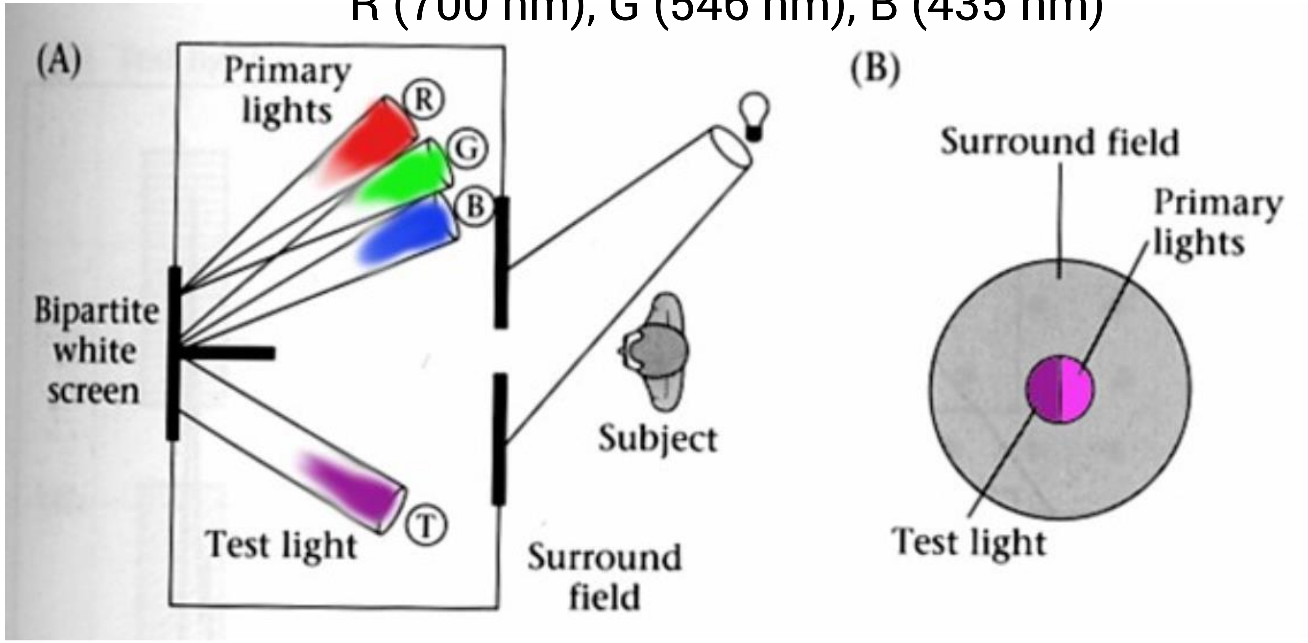


Color and Dynamic Range

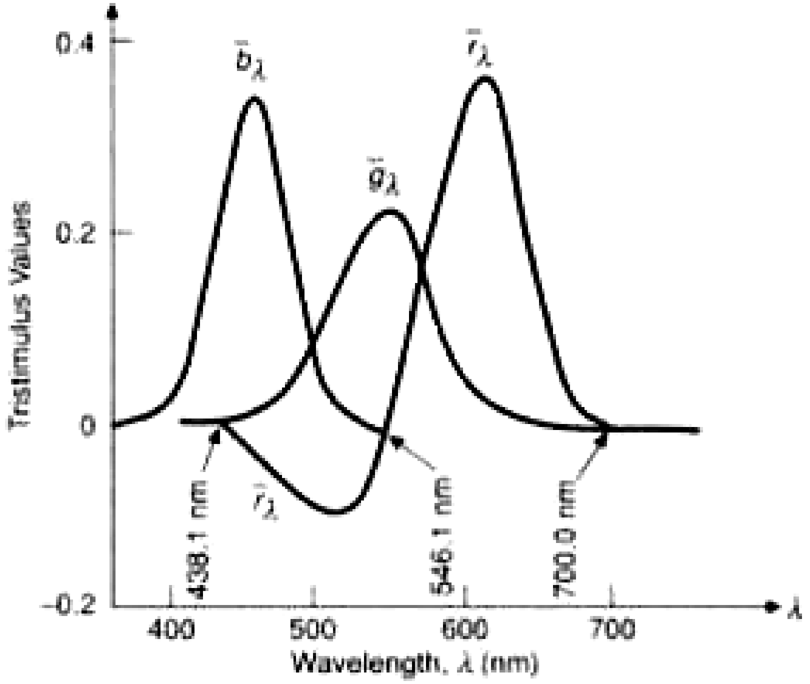
- Human Visual System
- **CIE Color Models**
- Reproducing Color
- Color Processing in Cameras

CIE RGB (1931)

R (700 nm), G (546 nm), B (435 nm)

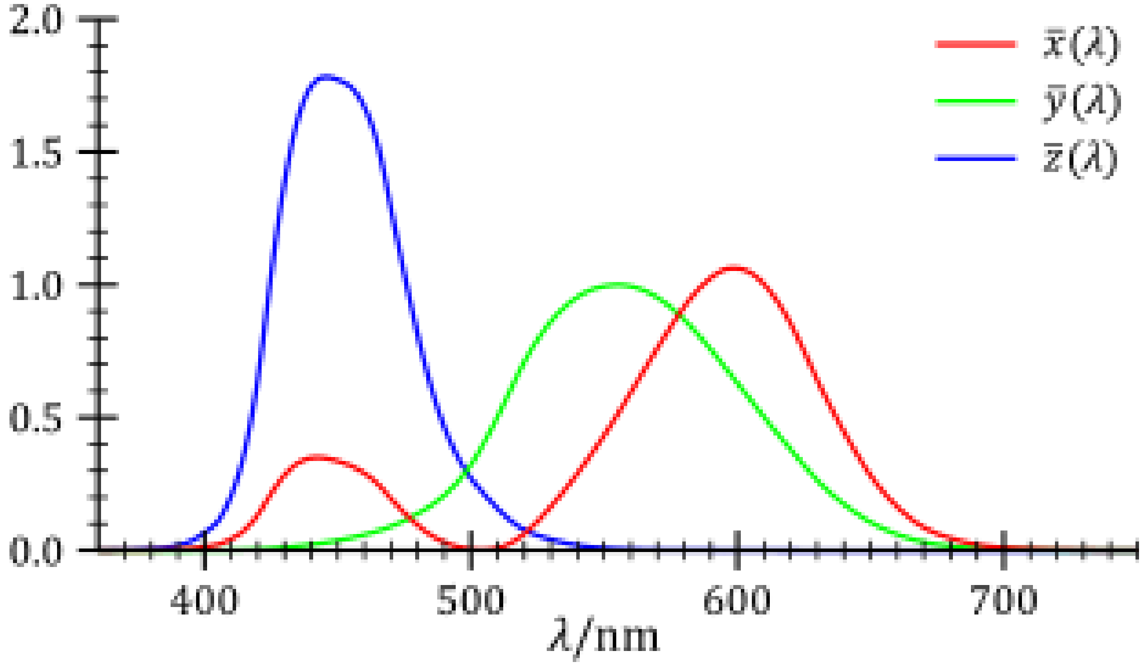


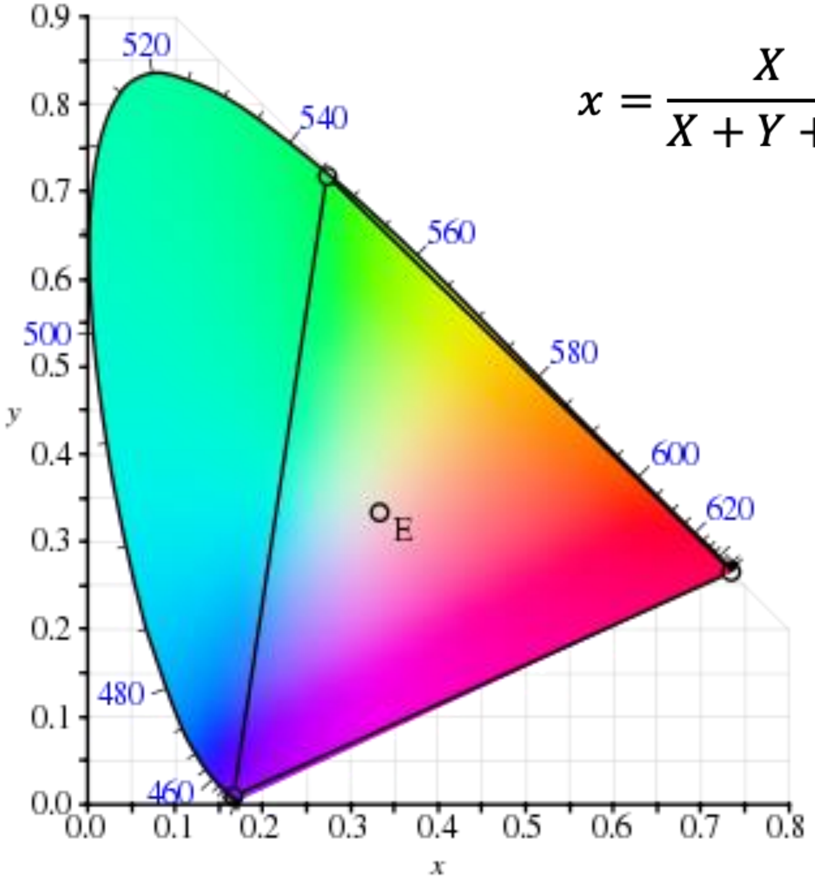
Foundations of Vision, by Brian Wandell, Sinauer Assoc., 1995



CIE XYZ and xy (1931)

$$X = \int_{\lambda=380nm}^{780nm} \phi(\lambda).\bar{x}(\lambda)d\lambda \quad \ldots$$





$$x = \frac{X}{X+Y+Z} \quad \ldots$$

Chromatic adaptation

Because the human visual system adapts to the surrounding light, CIE XYZ values are not absolute. They depend on the white point.

The CIE did their experiments using a D50 white point (sunny daylight).

When dealing with other white points, we need chromatic adaptation.

Von Kries transform

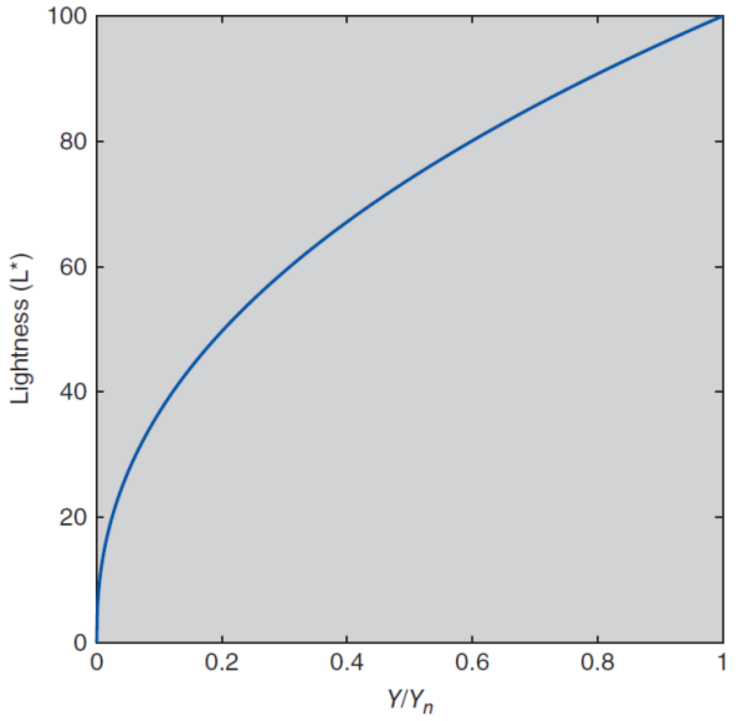
- Compute XYZ_{Illu} using $\bar{x}$, $\bar{y}$, $\bar{z}$, both for the color and the white point
- Convert everything to LSM (linear basis transform, i.e., matrix)
- Compute ratios $L_{D50}/L_{WhiteIllu}$, $M_{D50}/M_{WhiteIllu}$, $S_{D50}/S_{WhiteIllu}$
- Multiply LMS of color with these ratios
- Convert back to XYZ and obtain true XYZ_{D50}

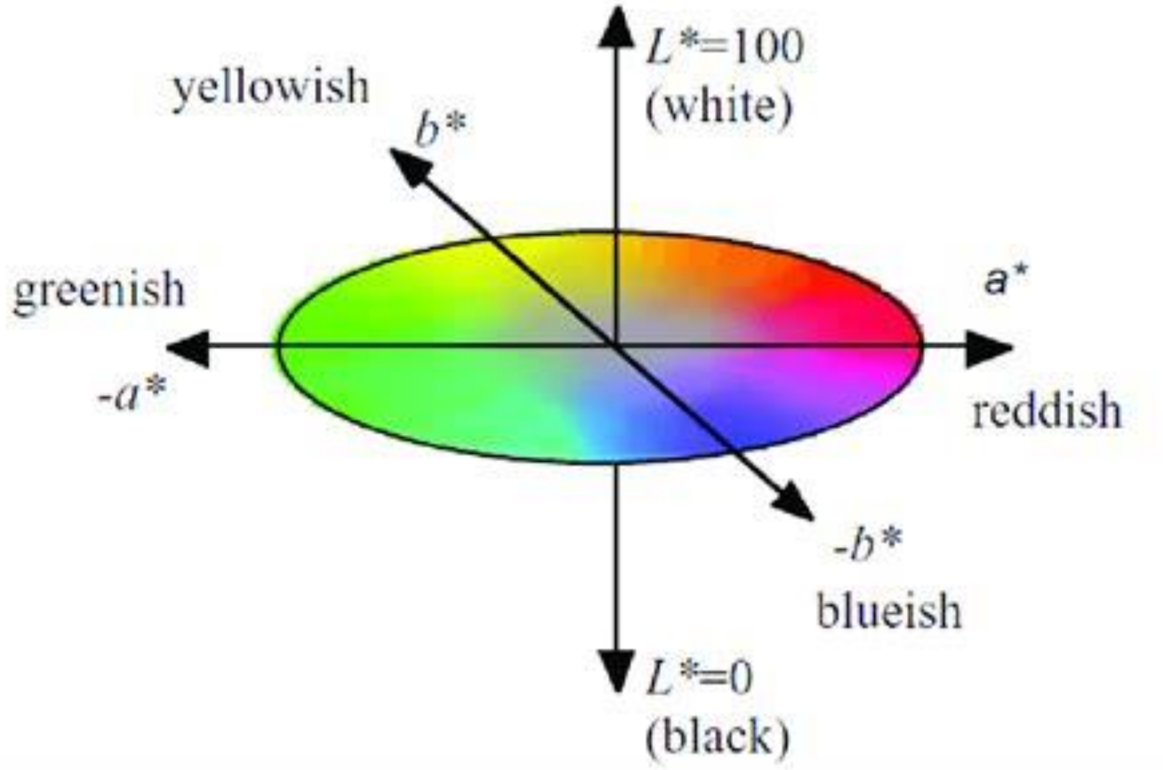
From color space to color appearance model: CIELAB and CIECAM

Goals:

- Compare colors, perceptually uniform color space
- Express colors through meaningful concepts such as brightness, saturation, hue, etc

Conversion includes non-linear steps because the human visual system is more sensitive to variations in the shadows.





Color and Dynamic Range

- Human Visual System
- CIE Color Models
- **Reproducing Color**
- Color Processing in Cameras

Color Display

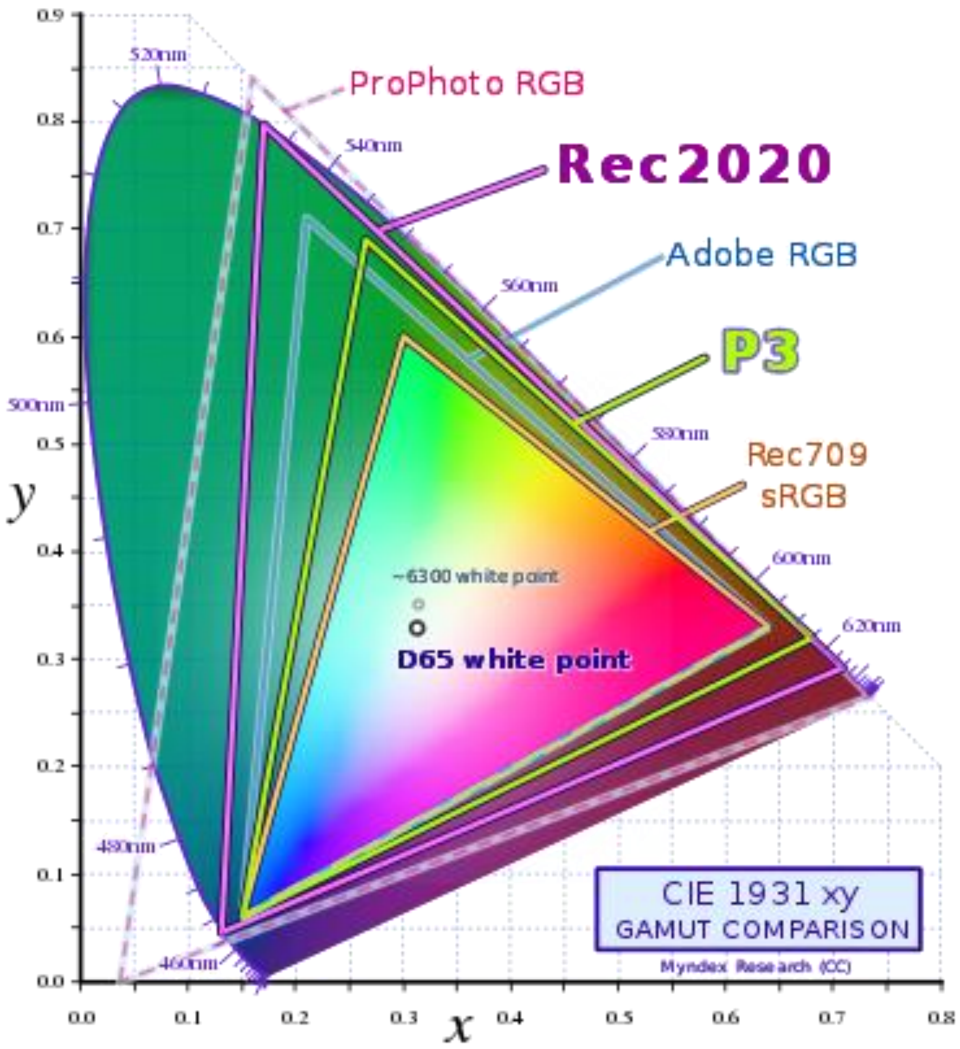
Thanks to trichromacy, we can reproduce any color with only 3 primary colors in our screen/projector: RGB!

The spectral power distributions are unimportant, we only care about the XYZ values of RGB primaries.

White point (what we obtain when R, G and B are at their maximum value) is important.

The "purer" the primaries are, the more colors we can reproduce: device gamut.

Trade-off between cost, efficiency and gamut.



Color Capture

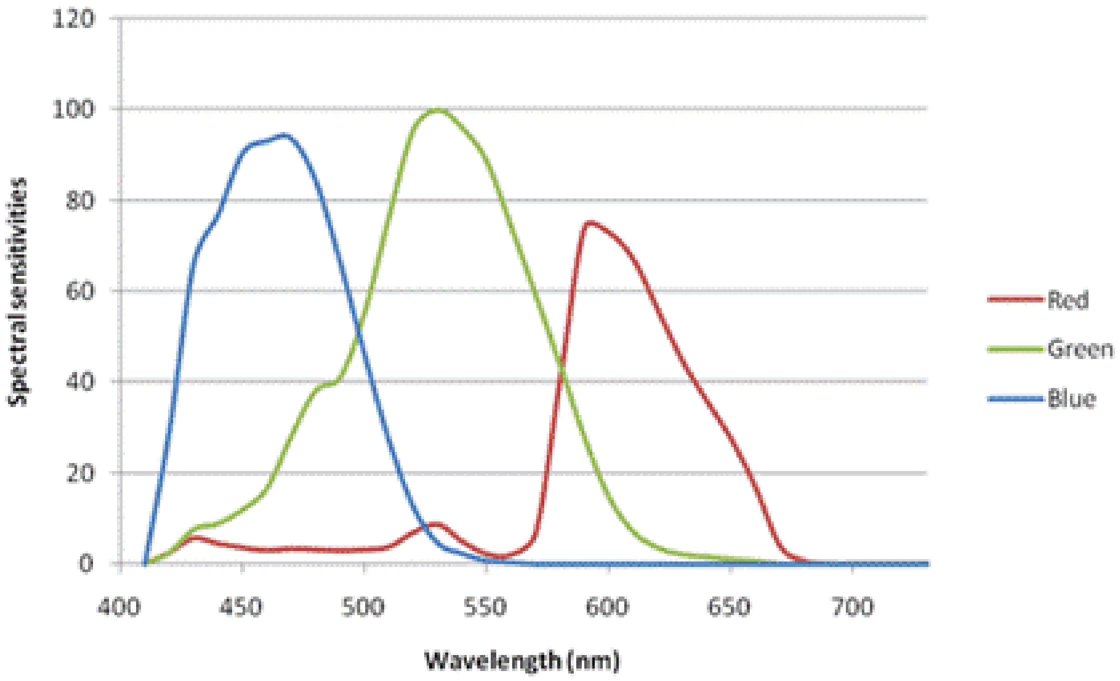
Capture devices don't have a gamut: every camera can capture every color (even if it's a black and white sensor). The challenge is to be able to distinguish between colors.

Most sensors use RGB Bayer filters.

Desired properties for their spectral response:

- Linear combination of $\ell(\lambda), m(\lambda), s(\lambda)$ to avoid Sensitivity Metamerism.
- Conversion from sensor RGB to display RGB should not increase noise (i.e., color matrix should have small singular values).
- Filter response should not be too narrow (don't waste precious photons).

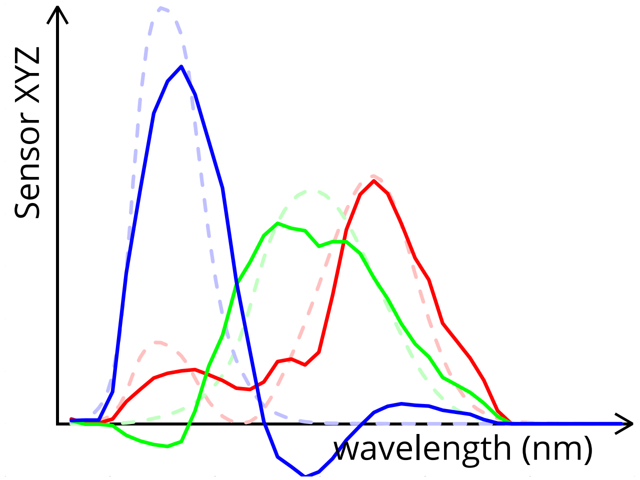
Overdetermined problem. Color filters can only be a trade-off between the above.



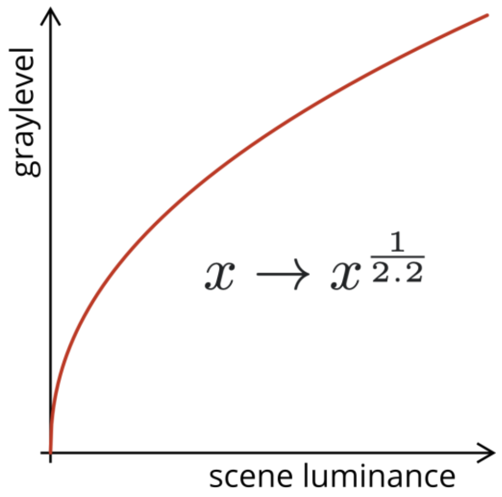
Converting from sensor native color to standard color (e.g. sRGB)

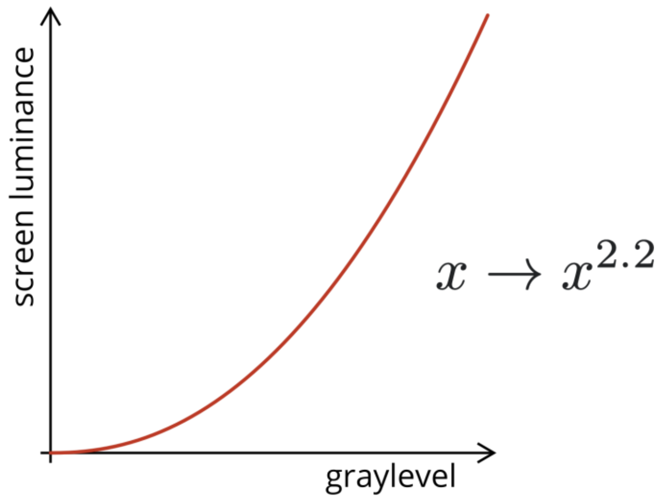
First step: primary conversion in linear domain

- Convert sensor RGB to XYZ (approximation)
- Convert XYZ to linear sRGB (exact)



Second step: apply transfer function (will be unapplied by the display).





Color and Dynamic Range

- Human Visual System
- CIE Color Models
- Reproducing Color
- **Color Processing in Cameras**

Alternative color models

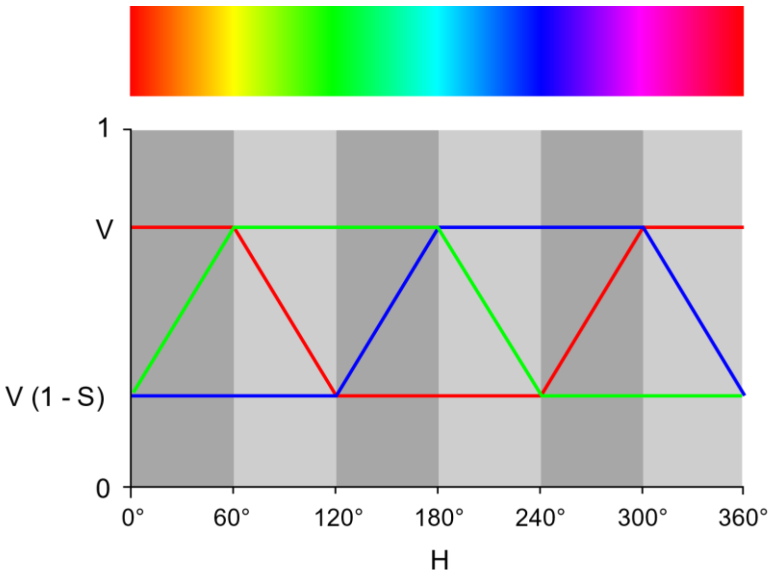
Introduce meaningful concepts such as hue and saturation.
Easier to convert from/to RGB than CIELAB. Often a good trade-off.

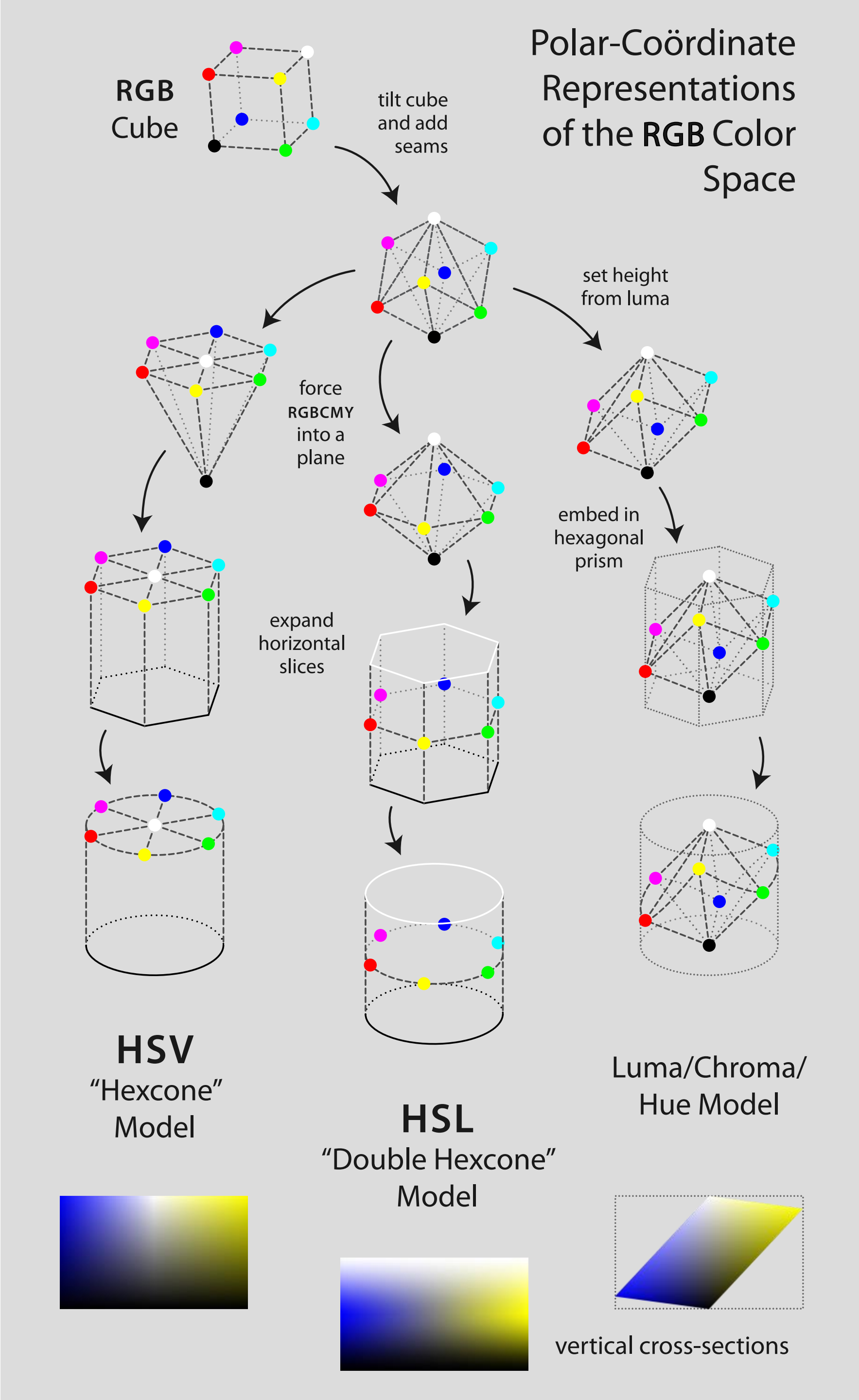
$$M = \max(R, G, B)$$
$$m = \min(R, G, B)$$
$$C = M - m$$

$$V = M$$
$$L = (M + m)/2$$
$$Y = 0.3R + 0.6G + 0.1B$$

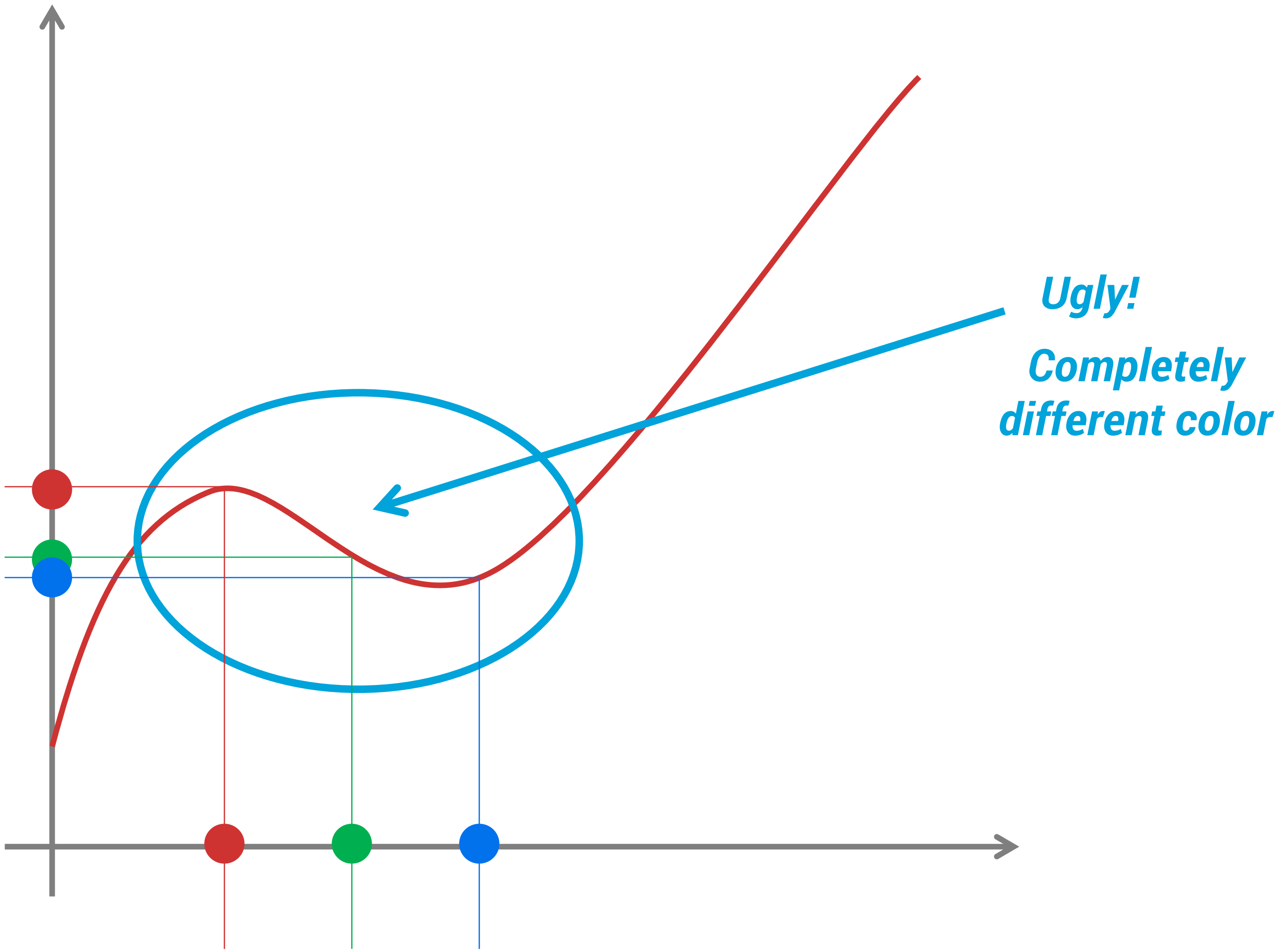
$$S = \frac{C}{V} \text{ or } \frac{C}{1-[2L-1|}$$

$$H = 60° \times \begin{cases} ? & \text{if } C = 0 \\ \frac{G-B}{C} \bmod 6 & \text{if } M = R \\ \frac{B-R}{C} + 2 & \text{if } M = G \\ \frac{R-G}{C} + 4 & \text{if } M = B \end{cases}$$



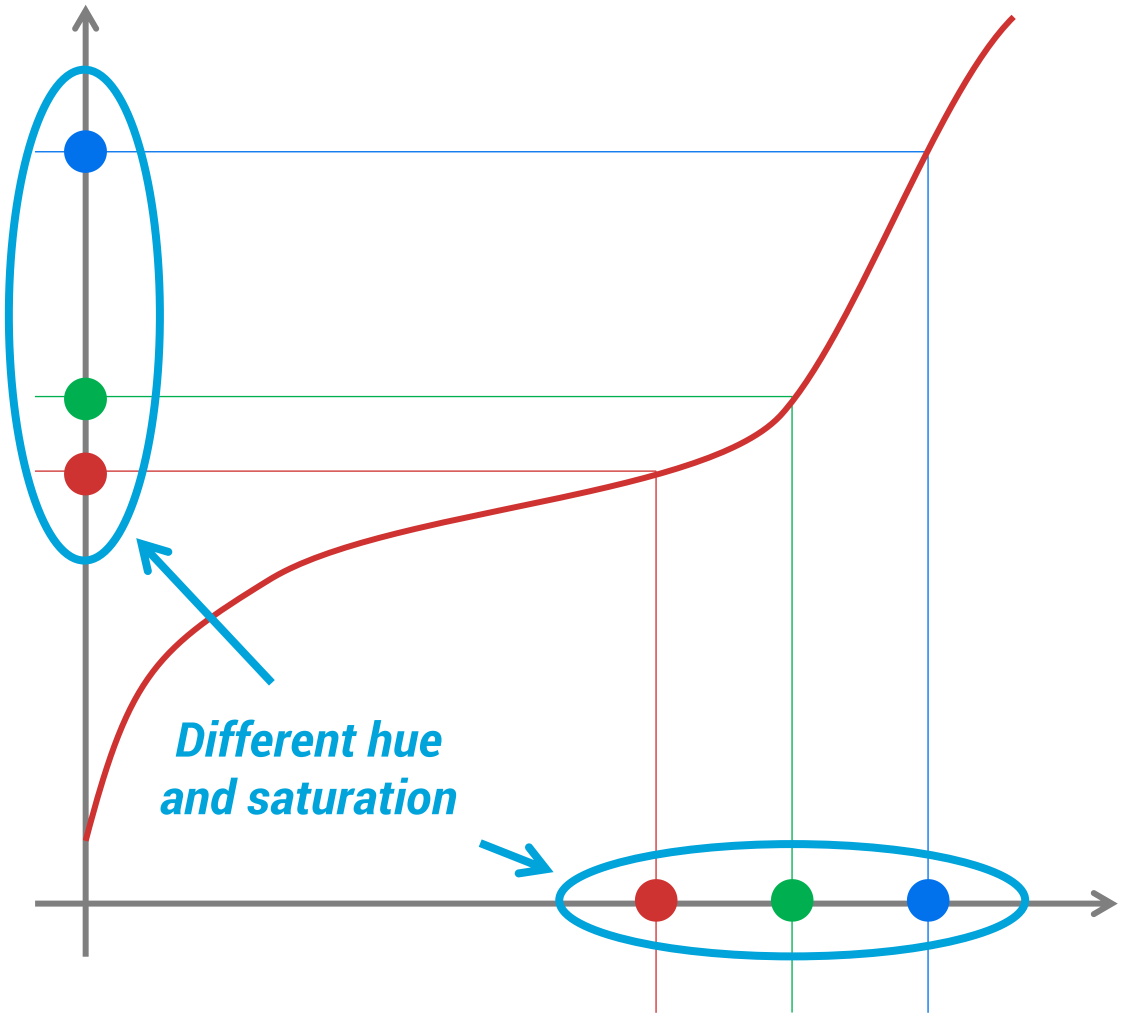


Don’t let the tone curve mess up your color



Tone curves should be monotonous

Don’t let the tone curve mess up your color



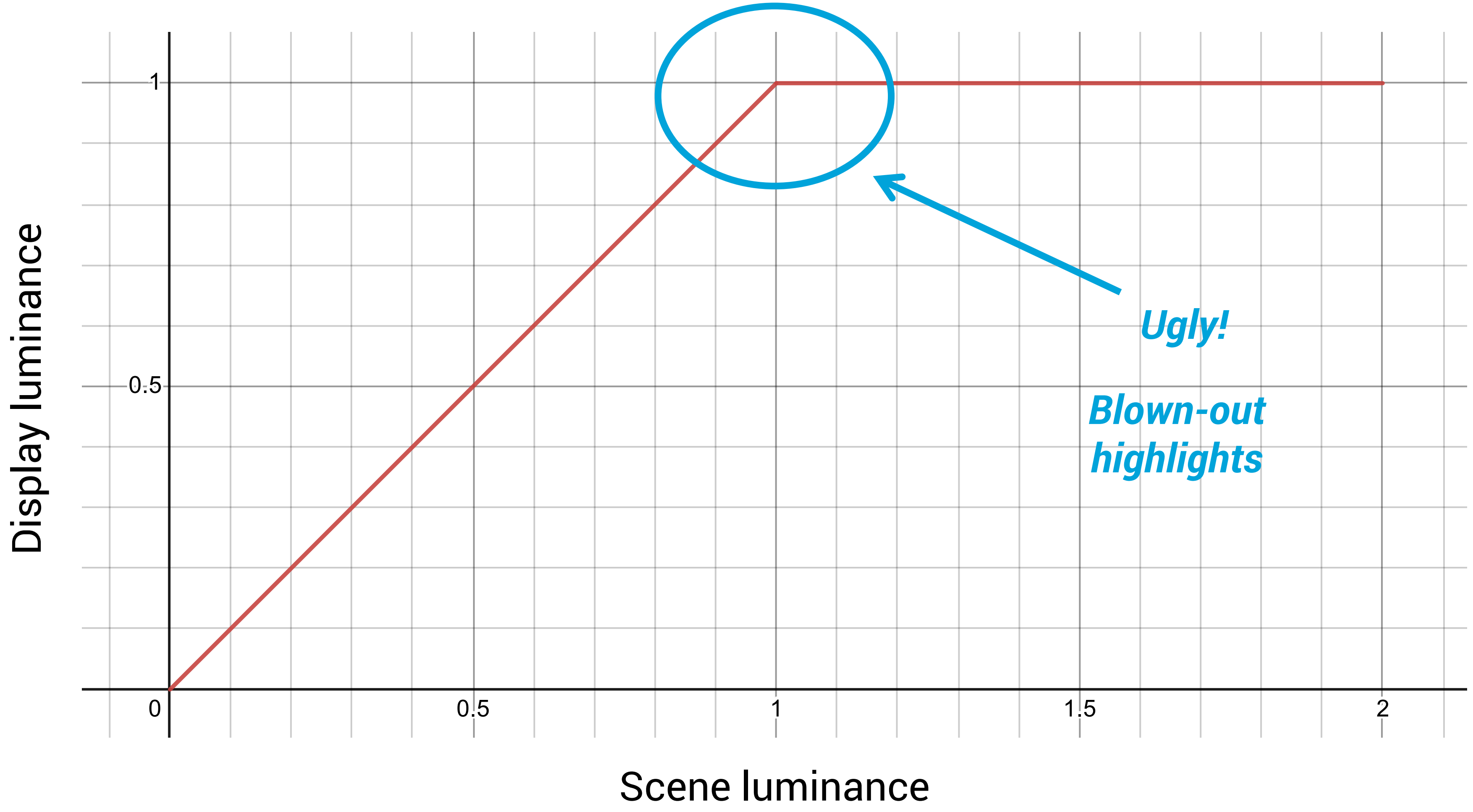
Tone curves (TC) should be monotonous.

Even when you apply the same TC to all channels, it changes not only the Y/L/V but also H and C/S. Can introduce H and C/S variations in Y/L/V gradients of constant H and C/S.

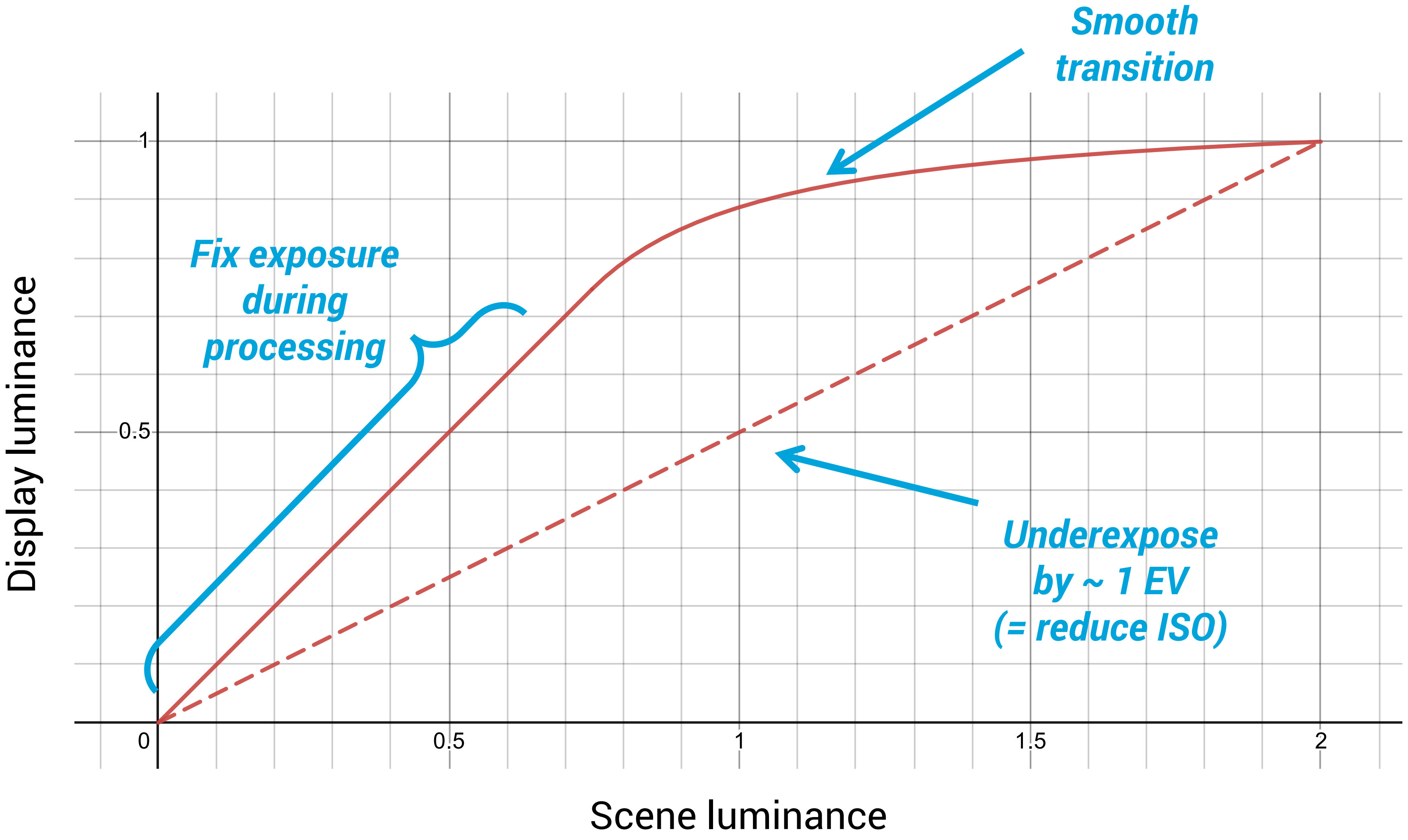
Enforce preservation of H and C/S: apply TC to max channel (V), compute ratio V_{out}/V_{in}, apply to all channels.

Enforce preservation of H: apply TC to max and min channels, compute H_{in}, then compute output median channel to preserve $H_{out}=H_{in}$.

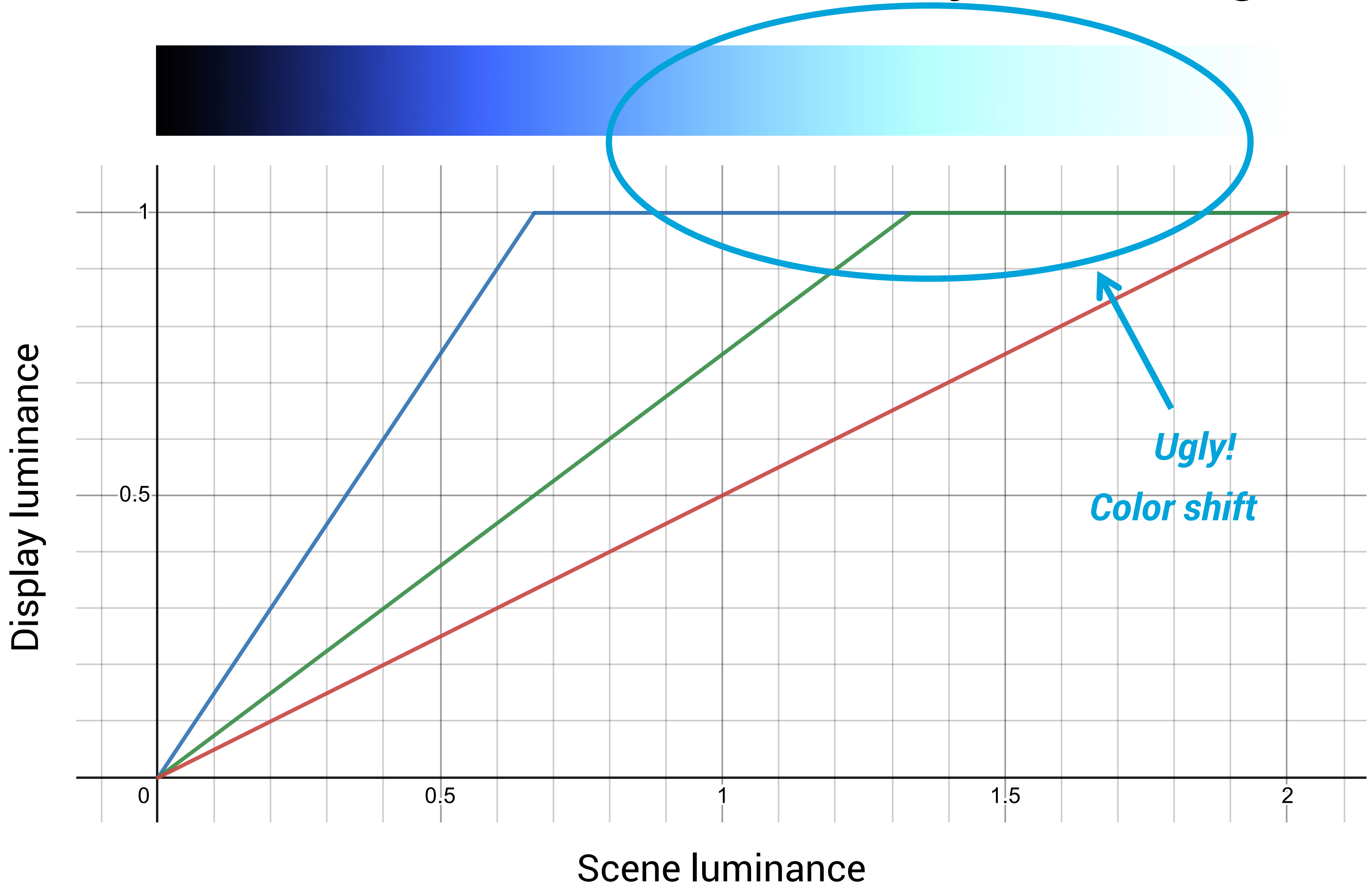
Attenuate nonlinear behavior due to limited dynamic range



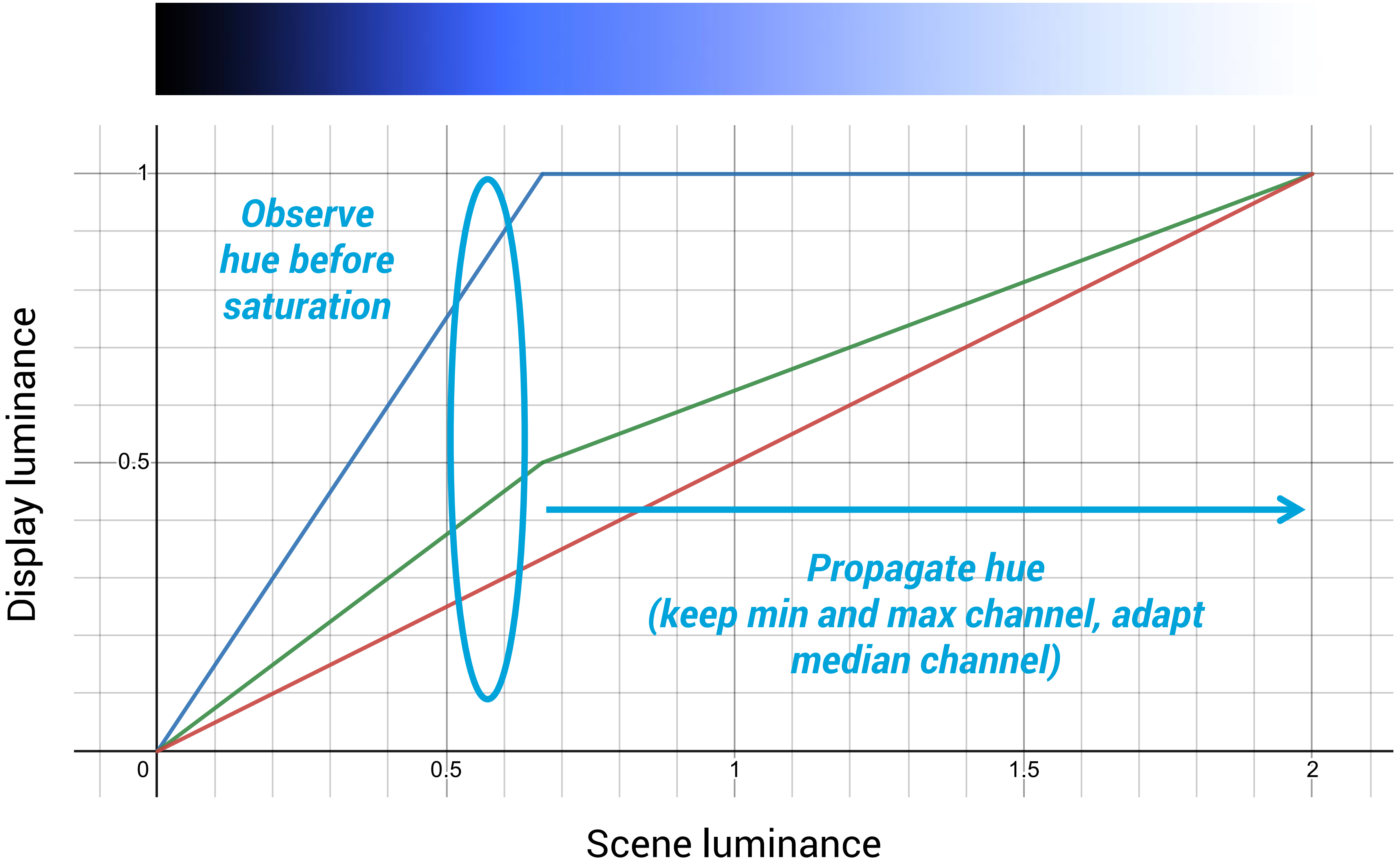
Attenuate nonlinear behavior due to limited dynamic range



Attenuate nonlinear behavior due to limited dynamic range



Attenuate nonlinear behavior due to limited dynamic range



Attenuate nonlinear behavior due to limited gamut

Color matrix Nikon Z9 (example)

$$\begin{pmatrix} 1.74 & -0.64 & -0.1 \\ -0.13 & 1.43 & -0.3 \\ 0.02 & -0.43 & 1.41 \end{pmatrix}$$

Coeffs > 1 on diagonal can create values > 1

Coeffs < 0 can create values < 0

Easiest solution: just clamp to [0, 1].
Leads to texture loss and colors shifts

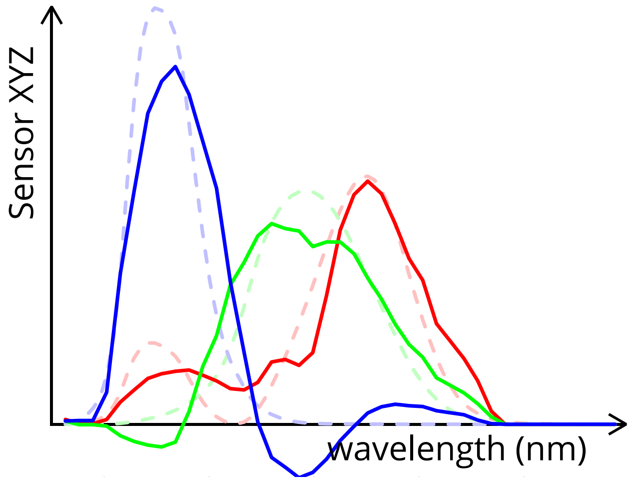
Use working color space with huge gamut, let user take care of the problem (via saturation sliders, applied locally).

Apply tricks that we have just seen: enforce hue preservation, attenuate large values to smooth the transition.

Most tone mapping operators can be adapted to do “gamut mapping”.

Introduce more degrees of freedom in the sensor calibration

Color matrix has only 6 degrees of freedom. But the spectral sensitivities of the color filters on the sensor are not simply a linear combination of those of the LMS (or XYZ) of the human visual system.
Consequently, the matrix cannot produce an exact solution.



Some approaches:

- Calibrate several matrices for different illuminants.
- Keep matrix model, but introduce more degrees of freedom by adding supplemental inputs: R^2, G^2, B^2, RG, RB, …, √RG, …
- Use look-up tables: RGB -> RGB, or H (“wavelength”) -> ΔHSV.
- Or any combination…

THANK YOU !

DXOMARK

DXO